Computational Details for HRV Analysis

We report mathematical details of some of the most used HRV features. This supplementary material complements Table 2 of the manuscript. Note, RR_j denotes the value of *j*'th RR interval and *N* is the total number of successive intervals

Time-domain HRV measures

Statistical measures.

- The simplest HRV feature is the mean value of the RR interval time-series, referred to as AVNN, and computed as:

$$AVNN = \frac{1}{N}\sum_{j=1}^{N} RR_j \quad \text{(Eq. 6)}$$

- The standard deviation of the RR intervals (SDNN) is:

$$SDNN = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}(RR_j - AVNN)^2} \quad \text{(Eq. 7)}$$

- The root mean square of successive differences (RMSSD) is:

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}(RR_{j+1} - RR_j)^2} \quad \text{(Eq. 8)}$$

- NN50, is another measure calculated from successive RR interval differences, computed as the number of successive intervals differing more than 50 ms and the corresponding relative amount pNN50 is:

$$pNN50 = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}\vartheta(\left|RR_{j+1} - RR_j\right| > 50\ ms)} \quad \text{(Eq. 9)}$$

where $\vartheta()$ is the Heaviside step function (i.e., the discontinuous function whose value is zero for negative argument and one for positive argument, formally $\vartheta(x) = 0$ if $x < 0$, otherwise $\vartheta(x) = 1$) and | | is the absolute value operator.

- Conceptually alike is the pNN10, computed on 10 ms:

$$pNN10 = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}\vartheta(\left|RR_{j+1} - RR_j\right| > 10\ ms)} \quad \text{(Eq. 10)}$$

In long-term analysis, two other parameters are usually computed.

- SDANN, the standard deviation of the averages of NN intervals in 5-minute segments:

$$SDANN = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}(RR_{5j} - RR_5)^2} \qquad \text{(Eq. 11)}$$

- SDNN IDX, the mean of the standard deviations of NN intervals in 5-minute segments:

$$SDNN\ IDX = \frac{1}{N}\sum_{j=1}^{N} SDNN_{5j} \qquad \text{(Eq. 12)}$$

Geometric measures.

- HRV triangular index is the integral of the density distribution divided by the maximum of the density distribution.

$$HRV\ IDX = \int \frac{D(t)dt}{Y} \qquad \text{(Eq. 13)}$$

- TINN (ms) - Triangular interpolation of NN interval histogram, which corresponds to the baseline width of the density distribution measured through triangular interpolation. To compute TINN, it is necessary to select values *N* and *M* and a multi-linear function q such that q(t) = 0 for t ≤ N and t ≥ M and q(X) = Y, and such that the integral in Eq.14:

$$\int_0^{+\infty}(D(t) - q(t))^2\, dt \qquad \text{(Eq. 14)}$$

is the minimum among all selections of all values *N* and *M.* The TINN measure, expressed in ms, is:

$$TINN = M\text{-}N \qquad \text{(Eq. 15)}$$

Frequency-domain HRV measures

The frequency domain HRV measures rely on the estimation of power spectral density (PSD), which could be computed with several methods This estimation can be made with non-parametric and parametric methods. Fast Fourier Transform (FFT) computation is the basis of the non-parametric PSD analysis, as summarized in Table 2.

Nonlinear HRV measures

Poincare Plot. The parameters of the Poincarè Plot SD_1 and SD_2 are usually computed according to the following formulae:

$$SD_1 = \frac{SDSD}{\sqrt{2}} \qquad \text{(Eq. 16)}$$

$$SD_2 = \sqrt{2SDNN^2 - \frac{1}{2}SDSD^2} \qquad \text{(Eq. 17)}$$

where SDSD is the standard deviation of the difference of RR interval time series.

Approximate Entropy (AppEn). The AppEn is usually computed according to the following algorithm. A series of vector of length m X_1, X_2, …, X_{N-m+1} is constructed from the RR intervals as follows:

$$X_i = [RR_i, RR_{i+1} \ldots RR_{i+m-1}]. \qquad \text{(Eq. 18)}$$

The distance $d[X_i, X_j]$ between vectors X_i and X_j is defined as the maximum absolute difference between their respective scalar components. For each vector X_i, the relative number of vectors X_j for which $d[X_i, X_j] \le r$, $C_i^m(r)$ is computed where *r* is referred as a tolerance value (Eq. 19).

$$C_i^m(r) = \frac{number\ of\ \{d[X_i, X_j] \le r\}}{N - m + 1} \quad \forall j \qquad \text{(Eq. 19)}$$

Then, the following index $\Phi^m(r)$ is computed by taking the natural logarithm of each $C_i^m(r)$ and averaging them over *i*.

$$\Phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \qquad \text{(Eq. 20)}$$

Finally, the approximate entropy is calculated as:

$$ApEn(m,r,N) = \Phi^m(r) - \Phi^{m+1}(r) \quad \text{(Eq. 21)}$$

Either *m*=1 or 2, and *r* between 0.1 and 0.2 times the SDNN, are suitable values to compute AppEn.

Sample Entropy (SampEn). SampEn computation is similar to AppEn, but with two key differences: (a) in the computation of $C_i^m(r)$ the comparison of the vector X(i) with itself is included in the count for AppEn, while this comparison is excluded for SampEn; (b) the logarithm is applied instead of subtraction in the final step. These changes aims to remove the bias in AppEn, as the count of the self-comparison in AppEn lower its value and the signals are interpreted as more regular than they are. The steps of SampEn computation are described as follows:

$$C_i^m(r) = \frac{number\ of\ \{d[X_i, X_j] \le r\}}{N-m+1} \quad \forall j \neq i \quad \text{(Eq. 22)}$$

$$\Phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \quad \text{(Eq. 23)}$$

$$SampEn(m,r,N) = \log \frac{\Phi^m(r)}{\Phi^{m+1}(r)} \quad \text{(Eq. 24)}$$

Correlation Dimension (CD). CD is computed similarly to AppEn. The reconstruction of the attractor is the first step to perform. That is, a series of vector of length *m* X_1, X_2, …, X_{N-m+1} is constructed from the RR intervals as follows:

$$X_i = [RR_i, RR_{i+\tau}, ..., RR_{i+\tau(m-1)}] \quad \text{(Eq. 25)}$$

where τ is the time delay and *m* is the embedding dimension. The second step is the estimation of Euclidean distances between each couple of vectors:

$$d[X_i, X_j] = \sqrt{\sum_{k=1}^{m} (X_i(k) - X_j(k))^2} \quad \text{(Eq. 26)}$$

Then, a function estimating the probability that two arbitrary points on the orbit are close than *r* is estimated. The correlation integral function is determined as:

$$C_m(r) = \frac{1}{N_m(N_m - 1)} \sum_{i}^{N_m} \sum_{j=1}^{N_m} \vartheta(r - d[X_i, X_j]) \quad \text{(Eq. 27)}$$

where $N_m = N - \tau(m-1)$ and ϑ is the Heaviside function.

The correlation dimension is defined as the following limit value:

$$CD(m) = \lim_{r \to 0} \lim_{N \to \infty} \frac{\log C^m(r)}{\log r} \quad \text{(Eq. 28)}$$

This limit value is approximated by the slope of the regression curve ($\log r$, $\log C^m(r)$). In HRV analysis the values of 1 and 10 are widely used value for τ and *m*, respectively.

Detrended Fluctuation Analysis (DFA). The DFA consists in the following steps:

a) The average $\overline{RR}$ of the RR interval series is calculated on all the *N* samples. The alternate component of the RR interval series, which is defined as the *RR* minus its average value $\overline{RR}$, is integrated:

$$y(k) = \sum_{j=1}^{k} (RR_j - \overline{RR}), k = 1,...., N \quad \text{(Eq. 29)}$$

b) The integrated series is divided into non-overlapping segments of equal length *n*. A least square line is fitted within each segment, representing the local trends with a broken line. This broken line is referred as $y_n(k)$, where *n* denotes the length of each segment.

c) The integrated time series is detrended as follows: $y(k) - y_n(k)$. The root-mean-square fluctuation of the detrended time series is computed according to the following:

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (y(k) - y_n(k))^2} \quad \text{(Eq. 30)}$$

d) The steps b) to d) are repeated for *n* from 4 to 64.

By representing the function *F(n)* in a log-log diagram, two parameters are computed: short-term fluctuations ($Alpha_1$) as the slope of the regression line relating $\log(F(n))$ to $\log(n)$ with *n* within

4-16; long-term fluctuations ($Alpha_2$) as the slope of the regression line relating $\log(F(n))$ to $\log(n)$ with n within 16-64.

Recurrence Plot. As in CD, vectors $X_i = (RR_i, RR_{i+\tau}, ..., RR_{i+(m-1)\tau})$, with $i=1, ..., K$, with $K = [N-(m-1)\tau]$, where m is the embedding dimension and τ is the embedding lag, are defined. The recurrence plot is a K-dimensional matrix of dots, in which one dot is placed if the Euclidean distance between X_i and X_j is lower than a threshold value r.

The following steps are suggested in order to obtain the recurrence plot:

a) A K-dimensional square matrix M_1 is calculated computing the Euclidean distances of each vector X_i from all the others.

b) A K-dimensional square matrix M_2 is calculated as the matrix whose elements $M_2(i,j)$ are defined as:

$$M_2(i,j) = \begin{cases} 1 & if \quad M_1(i,j) < r \\ 0 & if \quad M_1(i,j) > r \end{cases} \qquad \text{(Eq. 31)}$$

The recurrence plot is the representation of the matrix M_2 in which a dot is associated to one value, that is, an image in which black pixels correspond to ones and white pixels to zeros. M_1 is a symmetrical matrix as the distance between X_i and X_j is equal to the one between X_j and X_i and consequently, the recurrence plot is a symmetric image along the diagonal. Usually the following values of the parameters are chosen: $m = 10$; $\tau = 1$; $r = \sqrt{m} * SDRR$.

In the recurrence plot, lines are defined as series of diagonally adjacent black points with no white space. The length l of a line is the number of points constituting the line. Measures of the recurrence plot are widely computed: recurrence rate (REC) defined in Eq. 32; maximal length of lines (l_{max}); mean length of lines (l_{mean}); the determinism (DET) defined in Eq. 33; the Shannon entropy (ShEn) defined in Eq. 34.

$$REC = \frac{1}{K^2} \sum_{i=1}^{K} \sum_{j=1}^{K} M_2(i,j) \qquad \text{(Eq. 32)}$$

$$DET = \frac{\sum_{l=2}^{l_{max}} l * N_l}{\sum_{i=1}^{K}\sum_{j=1}^{K} M_2(i,j)} \quad \text{(Eq. 33)}$$

with N_l = number of lines of length l

$$ShEn = \sum_{l=l_{min}}^{l_{max}} n_l * \ln n_l \quad \text{(Eq. 34)}$$

with n_l = percentage of N_l over all the number of lines.